

David J. Chalmers

## A tudatos tapasztalat rejtélye\*

A tudatos tapasztalat egyszerre a legismerősebb és a legtitokzatosabb dolog a világon. Semmiről sincs közvetlenebb módon tudomásunk, mint a tudatról, de ezt a tudást rendkívül nehéz összeegyeztetni bármilyen más jellegű tudásunkkal. Mi okból létezik? Mire szolgál? Egyáltalán hogyan jöhet létre az agy idegi folyamataiból? Ezek a kérdések a tudomány egészének legizgalmasabbjai közé tartoznak.

Objektív szempontból az agy viszonylag jól megismerhető. Amikor ránézünk erre az oldalra, az információ-feldolgozás egész kis örvénye jön létre: fotonok csapódnak a retinánkra, elektromos jelzések továbbítódnak a látóidegen át és az agy különböző területei között, reakcióként pedig elmosolyodunk, megrökönyödve ráncoljuk a homlokunkat, esetleg megjegyzést teszünk. Mindig van azonban egy szubjektív szempont is. Az oldalnak tudatában vagyunk, amikor ránézünk, míg a képeket és szavakat, mint saját mentális életünk részeit, közvetlenül észleljük. Élénk észleleteink vannak a színes virágokról és a ragyogó égboltról. Eközben esetleg bizonyos érzelmeket élünk át és gondolatokat fogalmazunk meg. Az ilyen tapasztalatok együttese alkotja a tudatot: az elme belső, szubjektív életét.

Az agyat és az elmét tanulmányozó kutatók éveken át kerülték a tudat kérdését. Az uralkodó nézet az volt, hogy a tudomány, mely az objektivitáson alapul, nem fogadhat el egy oly mértékben szubjektív valamit, mint a tudat. A behaviorista mozgalom a pszichológiában, mely korábban a század folyamán uralkodó volt, a külső viselkedésre összpontosított, és helytelenítette a belső mentális folyamatokról való bármilyen beszédet. Később a kognitív tudomány megjelenése a fejen belüli folyamatokra terelte a figyelmet. A tudat mégis olyan tiltott terület maradt, amely csak az éjszakai, ital melletti beszélgetések körébe tartozik.

Az utóbbi években viszont egyre több idegtudós, pszichológus és filozófus veti el azt a gondolatot, mely szerint a tudat nem tanulmányozható, és próbálja kifürkészni annak titkait. Amint az egy ennyire új területtől elvárható, változatos és egymásnak ellentmondó, alapvető fogalmakat gyakran összeférhetetlen

\* [The Puzzle of Conscious Experience. *Scientific American*, Dec. 1995, 62–68. Az eredeti szöveghez hozzátartozik néhány szemlélető ábra, amelyeket nélkülözhetőknek ítélünk a gondolatmenet megértésének szempontjából. Ezeket lásd: <http://consc.net/papers/puzzle.html> – a szerk. megj.]

NB. Ahogyan a *Scientific American* esetében szokásos, a jelen szöveget is lényegesen lerövidítettük. A téma részletesebb tárgyalását lásd „Facing up to the Problem of Consciousness” című írásomban. [*Journal of Consciousness Studies*, 2(3)/1995, 200–219. Magyarul: David J. Chalmers: Szemközt a tudat problémájával. Ford. Sutyák Tibor. *Vulgo*, 2004/2, 14–35. – a szerk. megj.]

módon használó elméletek egész szövevénye jelent meg. A filozófiai okfejtés elengedhetetlen ahhoz, hogy mindezt megvilágíthassuk.

A nézetek miriádja ezen a területen belül mindent felölel a redukcionista elméletektől, melyek szerint a tudat megmagyarázható az idegtudomány és a pszichológia standard módszereivel, az úgynevezett „rejtély-elvűek” álláspontjáiig, akik azt állítják, hogy egyáltalán soha nem fogjuk megérteni a tudatot. Én úgy vélem, ha közelebbről szemügyre vesszük, akkor mindkét nézetről belátható, hogy hibás, és hogy az igazság valahol félúton található. A redukcionizmus ellen azzal fogok érvelni, hogy az idegtudomány eszközeivel a tudat nem magyarázható kimerítő módon, bár azok sokat nyújthatnak. A rejtély-elvűséggel szemben arra az álláspontra fogok helyezkedni, hogy a tudat egy újfajta elmélettel mégis megmagyarázható lehet. Egy ilyen elmélet még nem hozzáférhető minden részletében, de a gondos okfejtés és némi bölcs következtetés feltárhat valamennyit általában vett természetéből. Például: valószínűleg új alaptörvényeket fog majd létrehozni, és az információ fogalma központi szerepet játszhat ebben. Ez a halvány derengés azt sugallja, hogy a tudat elméletének megdöbbentő következményei lehetnek a világmindenségről, valamint önmagunkról vallott nézeteink vonatkozásában.

## A nehéz probléma

A kutatók a „tudat” kifejezést számos különböző módon használják. A kérdések tisztázása végett előbb külön kell választanunk a gyakran ugyanazon megnevezés alá besorolt problémákat. E célból hasznosnak látom megkülönböztetni a tudat „könnyű problémáit” a „nehéz problémától”. A könnyű problémák semmiképp nem banálisak – sőt, éppoly izgalmasak, mint bármi más a pszichológiában és a biológiában –, de a központi rejtély a nehéz problémához kapcsolódik. A tudat könnyű problémái közé tartoznak a következők: Hogyan képes az emberi szubjektum megkülönböztetni az érzéki ingereket, és megfelelően válaszolni ezekre? Hogyan egységesíti az agy a különböző forrásokból érkező információt, és hogyan használja a viselkedés ellenőrzésére? Hogyan lehetséges az, hogy a szubjektumok verbalizálni tudják belső állapotait? Ezek a kérdések mind a tudathoz kapcsolódnak ugyan, de mindegyikük a kognitív rendszer objektív mechanizmusaira vonatkozik. Következésképp minden okunk megvan feltételezni, hogy a kognitív pszichológia és az idegtudomány további kutatásai megválaszolják majd őket.

A nehéz probléma viszont azt veti fel, hogy az agyon belüli fizikai folyamatok hogyan hozhatják létre a szubjektív tapasztalatot. Ez a rejtély a gondolkodás és az észlelés belső vetületét foglalja magában, azt a módot, ahogyan a dolgok a szubjektum számára vannak. Amikor látunk, akkor vizuális érzeteink vannak, például az élénk kék szín érzete. Vagy gondoljunk egy távoli oboa megfoghatatlan hangzására, egy éles fájdalom kínjára, a boldogság könnyedségére, avagy egy gondolkodásba mélyedt pillanat meditatív minőségére. Ezek mind beletartoznak abba, amit tudatnak nevezek. Ezek azok a fenomének, melyek az elme valódi rejtélyét felvetik.

A különbségtétel szemléltetéséül tekintsünk egy Frank Jackson ausztrál filozófus által kieszelt gondolat kísérletet. Tegyük fel, hogy Mary, egy 23. századi idegtudós a világ legjobb szakembere a színlátásért felelős agyi folyamatok kérdésében. Mary azonban egész életét egy fekete-fehér szobában élte le, és sohasem látott színeket. Minden tudhatót tud az agyon belüli fizikai folyamatokról – az agy biológiájáról, struktúrájáról, funkciójáról. Ez a rálátás lehetővé teszi számára, hogy bármit felfogjon, ami a könnyű problémákhoz tartozik: hogyan különbözteti meg az agy az ingereket, hogyan egységesíti az információt, és hogyan számol be erről verbálisan. A színlátásra vonatkozó ismeretei által tudja, hogy a színek nevei hogyan függenek össze a fény spektrumának hullámhosszaival. De van valami a színlátást illetően, amit Mary nem ismer: az, hogy milyen egy színt észlelni, mondjuk a pirosat. Ebből következik, hogy vannak olyan dolgok a tudatos tapasztalatot illetően, amelyek az agy működésére vonatkozó adatokból nem vezethetők le.

Sőt, senki sem tudja, miért követi egyáltalán tudatos tapasztalat ezeket a fizikai folyamatokat. Miért van az, hogy amikor agyunk egy bizonyos hullámhosszú fényt dolgoz fel, a mélybibor tapasztalatát éljük át? Miért van egyáltalán tapasztalatunk? Nem végezhetette volna el éppolyan jól egy öntudatlan automata ugyanazokat a feladatokat? Ezek olyan kérdések, melyekre valamilyen tudatelméletől várnánk választ.

Nem tagadom, hogy a tudat az agyból lép elő. Tudjuk például, hogy a látás szubjektív tapasztalata szorosan összefügg a látóközponthoz végbemenő folyamatokkal. Csakhogy maga az összefüggés az, ami zavarbaejtő. Érdekes módon a szubjektív tapasztalat egy fizikai folyamatból tűnik előlépni. Csakhogy fogalmunk sincs, hogyan.

## Elégséges-e az idegtudomány?

Látván a tudat tanulmányozásának jelenlegi sikereit az idegtudományban és a pszichológiában, azt hihetnénk, a titok megfejtése elkezdődött. Közelebről megvizsgálva viszont kiderül, hogy szinte mindegyik kortárs kutatás csak a tudat könnyű problémáit veti föl. A redukcionista felfogás magabiztossága a könnyű problémák tárgyában történt előrehaladásból táplálkozik, de mindez lényegtelen a nehéz probléma vonatkozásában.

Vegyük Francis Crick, a San Diego-i Salk Institute for Biological Studies, valamint Christof Koch, a California Institute of Technology neurobiológusa által felállított hipotéziseket. Ők azt vetik föl, hogy a tudat az agykéreg bizonyos rezgéseiből jöhet létre, melyek szinkronizálódnak a neuronok másodpercenkénti 40 kisülése folyamán. Crick és Koch úgy gondolják, a jelenség megmagyarázhatja, hogy egy egyedi észlelt tárgy különböző tulajdonságai (például a színe és a formája), melyeket az agy különböző részei dolgoznak fel, hogyan állhatnak össze egyetlen koherens egészé. Ezen elmélet szerint két információ-részlet pontosan akkor áll össze, amikor szinkronizált idegi kisülések reprezentálják őket.

Elképzelhető, hogy a hipotézis megoldhatja a könnyű problémák egyikét arra nézve, hogy az információ hogyan áll össze az agyban. De miért kellene az

összehangolt rezgéseknek látási tapasztalathoz vezetniük, függetlenül attól, hogy mennyire egységesítettek? Ez a kérdés a nehéz problémát foglalja magában, amelyre nézve az elmélet semmit sem tud nyújtani. Sőt, Crick és Koch agnosztikusak abban a kérdésben, hogy a nehéz probléma tudományosan egyáltalán megoldható-e.

Ugyanazt a kritikát alkalmazhatnánk a kortárs tudat kutatás legnagyobb részére. 1991-es *Consciousness Explained* című könyvében Daniel C. Dennett filozófus egy bonyolult elméletet mutatott be arra nézve, hogy nagy számú független agyi folyamat hogyan állhat össze egy észlelt eseményre adott egyetlen koherens válaszá. Az elmélet nagyban hozzájárulhat annak magyarázatához, hogyan számolunk be verbálisan belső állapotainkról, de igen keveset mond arról, miért kellene szubjektív élménynek meghúzódnia a verbális beszámoló mögött. Más redukcionista elméletekhez hasonlóan a Dennett-é is a könnyű problémák elmélete.

A döntő közös jellemvonás a könnyű problémák között az, hogy mind arra vonatkoznak, ahogyan egy kognitív vagy viselkedési funkció működik. Végső soron mindegyik arra kérdez rá, hogy az agy hogyan végez el valamilyen feladatot – hogyan különböztet meg ingereket, hogyan egységesíti az információt, hogyan számol be erről, és így tovább. Amint a neurobiológia meghatározza majd a megfelelő idegi mechanizmusokat, megmagyarázván a funkciók működését, a könnyű problémák megoldódnak. A tudat nehéz problémája viszont messze túlmege azokon a kérdéseken, hogy a funkciók hogyan működnek. Még ha minden olyan viselkedési és kognitív funkciót meg is magyaráznánk, amely a tudattal összefügg, akkor is maradna egy további rejtély: miért kíséri e funkciók elvégzését tudatos tapasztalat? Ez a külön talány az, ami a nehéz problémát nehezzé teszi.

## A magyarázati szakadék

Vannak, akik felvetik, hogy a nehéz probléma megoldásához új fizikai magyarázati eszközökre van szükség: nem-lineáris dinamikára, mondjuk, vagy új idegtudományi felfedezésekre, vagy kvantummechanikára. De ezek az elképzelések ugyanabba a nehézségbe ütköznek. Tekintsük az Arizonai, illetve az Oxfordi Egyetemen dolgozó Stuart R. Hameroff, valamint Roger Penrose ajánlatát. Ők úgy tartják, hogy a tudat kvantumfizikai folyamatokból jön létre bizonyos mikrosövecskékben, melyek az idegsejteken belüli fehérje-szerkezetek. Lehetséges (ha nem is túl valószínű), hogy egy ilyen hipotézis elvezet majd annak magyarázatához, hogy az agy hogyan hoz döntéseket, vagy éppen hogyan bizonyít matematikai tételeket, ahogy azt Hameroff és Penrose felvetik. De még ha ez meg is történik, az elmélet semmit sem fog mondani arról, hogy ezek a folyamatok hogyan hozhatnak létre tudatos tapasztalatot. Sőt, ugyanaz a probléma bármely, kizárólag a fizikai feldolgozáson alapuló tudatelmélet kapcsán felmerül.

A baj az, hogy a fizikai elméletek a legalkalmasabbak arra, hogy megmagyarázzák, egy rendszer miért rendelkezik bizonyos fizikai struktúrával, és hogyan végez bizonyos működéseket. A legtöbb tudományos probléma ilyen

formában jelenik meg; ahhoz, hogy például az életet magyarázzuk, azt kell megmagyaráznunk, hogyan képes szaporodni, alkalmazkodni és anyagcserét végezni egy fizikai rendszer. De a tudat teljesen másfajta probléma, mivel túlmegy a struktúra és a működés magyarázatán.

Természetes, hogy az idegtudomány nem nélkülözhető a tudat tanulmányozásában. Egyrészt lehetséges, hogy feltárja majd a tudat idegi korrelátumának a természetét – azt az agyi folyamatot, amely a legközelebből összefügg a tudatos tapasztalattal. Talán még részletes megfelelést is kimutat majd bizonyos agyi folyamatok és a velük összefüggő tapasztalat-összetevők között. De amíg megtudjuk, hogy ezek a folyamatok miért hoznak létre tudatos tapasztalatot egyáltalán, nem fogjuk áthidalni azt, amit Joseph Levine filozófus a fizikai folyamatok és a tudat közötti magyarázati szakadékknak nevezett. A szakadék átugrásához egy újfajta elméletre lesz szükség.

## Mindenek igaz elmélete

Az alternatíva keresése közben kulcsfontosságú megtenni azt az észrevételt, hogy a tudomány területén nem minden entitás magyarázható további, alapvetőbb entitásokkal. A fizikában például a tér-időt, a tömeget és a töltést (többek között) a világ alapvető tulajdonságainak tekintik, mivel nem vezethetők vissza semmi egyszerűbbre. A visszavezetés lehetetlensége mellett mégis részletes és hasznos elméletek állapítják meg viszonyaikat bizonyos alaptörvények értelmében. E tulajdonságok és törvények együtt nagyszámú komplex és nehezen megfogható jelenséget magyaráznak meg.

Széles körben elterjedt az a nézet, hogy a fizika az univerzum alapvető tulajdonságai és törvényei teljes jegyzékével rendelkezik. Ahogy azt Steven Weinberg *Dreams of a Final Theory* című 1992-es könyvében megfogalmazza, a fizika célja létrehozni „mindenek elméletét”, amelyből mindent le lehet vezetni, ami az univerzumból egyáltalán tudható. Weinberg azonban elismeri, hogy a tudat problémát jelent. A fizikai elmélet magyarázóerejének ellenére a tudat létezése mégsem tűnik levezethetőnek a fizika törvényeiből. A fizika védelmében úgy érvel, hogy az megmagyarázhatja az úgynevezett objektív tudatkorrelátumokat (azaz az idegi tudatkorrelátumokat), de ez még nyilván nem egyenlő a tudatnak magának a magyarázatával. Ha a tudat létezése nem vezethető le a fizika törvényeiből, akkor a fizikai elmélet nem mindenek igaz elmélete. A végső elméletnek tehát egy kiegészítő alapösszetevőt kell tartalmaznia.

Ebből a célból azt ajánlom, hogy tekintsük a tudatos tapasztalatot olyan alapvető tulajdonságnak, amely semmi alapvetőbbre nem vezethető vissza. Az elképzelés első látásra furcsának tűnhet, de úgy tűnik, a következetesség szempontjából mégis szükségzerű. A 19. században kiderült, hogy az elektromágneses jelenségeket nem lehet az addig ismert elvek alapján magyarázni. Következésképp a tudósok bevezették az elektromágneses töltést mint alapvető entitást, és az ezzel kapcsolatos alaptörvényeket kezdték tanulmányozni. A tudatra is hasonló gondolatmenetet kell alkalmaznunk. Ha a létező alapelméletek nem tudják átfogni, akkor valami újra van szükség.

Ahol van valamilyen alaptulajdonság, ott vannak alaptörvények is. Ebben az esetben a törvényeknek a tapasztalatot kell összekötniük a fizikai elmélet elemeivel. Ezek a törvények szinte biztosan nem fogják keresztezni a fizikai világ törvényeit; úgy tűnik, hogy ez utóbbiak már önmagukban zárt rendszert alkotnak. A törvények inkább hídként fognak szolgálni, megmutatván, hogy a tapasztalat hogyan függ az őt megalapozó fizikai folyamatoktól. Ez a híd az, ami a magyarázati szakadékot átíveli majd.

A teljes elméletnek tehát két összetevője lesz: a fizikai törvények, amelyek a fizikai rendszerek viselkedéséről számolnak be a végtelenül kicsitől a végtelenül nagyig, és azok, amelyeket pszichofizikai törvényeknek nevezhetnénk, és arról számolnak be, hogy e rendszerek hogyan függenek össze a tudatos tapasztalattal. Ez a két összetevő mindenek igaz elméletét fogja létrehozni.

## Az elmélet nyomában

Feltételezvé, hogy léteznek, hogyan fedezhetnénk fel ezeket a pszichofizikai törvényeket? A kutatás legnagyobb akadálya az adathiány lenne. Leírásom szerint a tudat szubjektív, tehát nincs közvetlen mód rá, hogy másoknak megjelenítsük. De ez a nehézség csak akadály, nem zsákutca. Először is, mind egyikünk hozzáférhet saját élményeihez, s ez a gazdag leletforrás felhasználható elméletek megfogalmazására. Emellett közvetett információkra, például az alanyok saját élményeiről adott leírására is támaszkodhatunk. A filozófiai érvek és gondolat kísérletek szintén szerepet játszhatnak. Az ilyen módszereknek megvannak ugyan a maguk korlátai, de kezdetnek több mint elégséges anyaggal szolgálnak.

Ezek az elméletek nem lesznek meggyőző módon kipróbálhatók, így elkerülhetetlenül spekulatívabbaknak kell majd lenniük a hagyományosabb tudományos diszciplínáknál. Ennek ellenére nincs rá ok, miért ne lennének rákényszerítve, hogy első személyű tapasztalatainkat az alanyok beszámolóinak adataival együtt igen szabatosan leírják. Ha találunk olyan elméletet, amely egy másik, azonos bonyolultsági fokú elméletnél jobban ráillik az adatokra, jó okunk van arra, hogy elfogadjuk azt. Pillanatnyilag egyetlen olyan elméletünk sincs, amely ráillene az adatokra, úgyhogy korai a kipróbálhatóság miatt aggódnunk.

Kezdhethetnénk azzal, hogy magas szintű áthidaló törvényeket keresünk, melyek hétköznapi szinten kötnék össze a fizikai folyamatokat a tapasztalattal. Egy ilyen törvény alapjaiban körvonalazható lehet azon megfigyelés alapján, hogy amikor valaminek a tudatában vagyunk, olyankor általában képesek vagyunk annak értelmében cselekedni és beszélni róla – ezek pedig objektív, fizikai funkciók. Fordítva, amikor valamely információ közvetlenül hozzáférhető a cselekvés és a beszéd számára, olyankor az általában tudatos. Így tehát a tudat jól korrelációba állítható azzal, amit „tudatosságnak” nevezünk: azzal a folyamattal, amely által az agyban lévő információ általánosan hozzáférhetővé válik olyan motorikus folyamatok számára, mint a beszéd vagy a testi cselekvés.

Lehet, hogy a fogalom túl köznapinak tűnik, ám a tudatosság, ahogyan itt meghatározzuk, objektív és fizikai, míg a tudat nem az. A tudatosság fogal-

mának még finomításra van szüksége ahhoz, hogy a beszéd képességével nem rendelkező állatokra és a csecsemőkre is kiterjeszthessük, viszont, legalábbis az ismerős esetekben, láthatóvá válik a pszichofizikai törvény hozzávetőleges körvonala: ahol tudatosság van, ott tudat is van, és fordítva.

Hogy egy lépéssel továbbvigyük ezt a gondolatmenetet, tekintsük a tudatos tapasztalatban fennálló struktúrát. Egy látómező tapasztalata például egy folyamatosan változó szín-, alak- és minta-mozaik, és mint ilyen, részletes mértani struktúrával rendelkezik. A tény, hogy ezt a struktúrát le tudjuk írni, meg tudjuk ragadni különböző elemeit, és más, tőle függő cselekvéseket tudunk végrehajtani, azt sugallja, hogy a struktúra közvetlenül megfelel azon információ szerkezetének, amely az agyban a tudatosság idegi folyamatainak keresztül hozzáférhetővé válik.

Ehhez hasonlóan, szintapasztalatunk belső háromdimenziós szerkezettel bír, mely az agy látóközpontján belüli információs folyamatok struktúrájában tükröződik. Ezt a struktúrát tükrözik a művészek által használt színgorogok és színtérképek. A színeket rendszeres formába szervezik – vöröstől zöldig az egyik tengelyen, kéktől sárgáig a másikon, és fehértől feketéig a harmadikon. A színgorogon egymáshoz közel eső színeket hasonlókként észleljük. Igen valószínű, hogy ezek hasonló érzéki képzeteknek felelnek meg az agyban is, mint egy olyan komplex háromdimenziós idegsejtközi kódrendszer részei, amelyet még nem értünk eléggé. Az alapul szolgáló fogalmat átalakíthatjuk strukturális koherencia-elvvé: az információ szerkezete a tudatosságban a tudatos tapasztalat szerkezetét tükrözi, és megfordítva.

Egy másik jelölt a pszichofizikai törvény címre a szerveződési invariancia elve. Ez azt állítja, hogy az ugyanolyan absztrakt szerveződéssel bíró fizikai rendszerek ugyanolyan tudatos tapasztalatra fognak vezetni, függetlenül attól, hogy miből vannak. Például ha az idegsejtjeink közötti pontos kölcsönhatásokat szilícium-*chipek*kel le tudnánk másolni, akkor ugyanaz a tudatos tapasztalat jelenne meg. Az elgondolás valamelyest vitatható, de szerintem igen jól alátámasztható olyan gondolat kísérletekkel, amelyek a neuronok szilícium-*chipek*kel való fokozatos helyettesítését írják le. Figyelemre méltó implikációja ezen elképzelésnek, hogy egyszer a tudat gépekben is létrehozható lesz.

## Fizikai és tapasztalati információ

A tudat elméletének alapvető célkitűzése egy egyszerű és elegáns alaptörvény-csomag, a fizika alaptörvényeinek analógiájára. A fent leírt elvek viszont valószínűleg nem alapvetőek. Inkább magas szintű pszichofizikai elveknek tűnnek, amelyek a fizika olyan makroszkopikus elveivel analógok, mint amilyenek a termodinamika vagy a kinematika elvei. De melyek ezek további alapelvei? Senki sem tudja, de engedjék meg, hogy mégis spekuláljak.

Feltételezem, hogy az alapvető pszichofizikai törvényekben központi szerepe van az információ fogalmának. Az információ elvont fogalma, ahogy azt az 1940-es években a Massachusetts Institute of Technology-n dolgozó Claude E. Shannon megfogalmazta, a következő: különálló állapotok, amelyek között

a hasonlóságok és különbségek elemi szerkezete áll fenn. Példaként gondoljunk egy 10 bites bináris kódra mint információs állapotokra. Ilyen információs állapotokat a fizikai világ is magában foglalhat. Ez olyankor áll fenn, ha azok olyan fizikai állapotoknak felelnek meg (például elektromos feszültségeknek), amelyek különbségeit valamely útvonalon, például egy telefonvonalon továbbítani lehet.

A tudatos tapasztalat információkat foglal magában. A látómezőben megjelenő színfoltok mintázatát értelmezhetjük például a képernyő-pixelek mintájára. Érdekes módon az derül ki, hogy ugyanazokat az információs állapotokat találjuk a tudatos tapasztalatban, mint amelyek az agy fizikai folyamatait alapozzák meg. A szinterek háromdimenziós kódolása például arra mutat, hogy a valamely szín tapasztalásában megjelenő információs állapot közvetlenül megfelel egy, az agyban lévő információs állapotnak. A két állapotot tekintetnénk akár ugyanazon információs állapot két különböző vetületének, amely egyidejűleg foglaltatik benne a fizikai feldolgozásban és a tudatos tapasztalatban is.

Innen értelemszerűen adódik egy hipotézis. Lehetséges, hogy az információnak két elemi vetülete van: egy fizikai és egy tapasztalati. Ez a hipotézis alapvetéssel bírhat a fizikai folyamatok és a tapasztalat összefüggésére nézve. Ahol tudatos tapasztalat található, ott az egy olyan információs állapot egyik vetületé-ként létezik, amelynek másik vetületét az agy fizikai folyamatai foglalják magukban. Ez az indítvány még kiegészítésekre szorul ahhoz, hogy kielégítő elmélet váljék belőle. Szépen ráillik viszont a korábban említett elvekre – arra például, hogy az azonos szerveződésű rendszerek ugyanazt az információt foglalják magukban –, és tudatos tapasztalatunk számos tulajdonságát magyarázhatja meg.

Ez az elképzelés jónéhány másikkal is összefér, mint például John A. Wheeler találmányával, mely szerint az információ alapvető a világegyetem fizikájára nézve. Lehetséges, hogy a fizika törvényeit végső megközelítésben információs fogalmakban fogalmazzák majd meg, amely esetben a fizikai és a pszichofizikai törvények felépítménye kielégítő módon egybevágná. Még az is lehetséges, hogy a fizika elmélete és a tudatelmélet végül egyetlen átfogóbb információ-elméletté lenne egységesíthető.

Problémát jelenthet az információ mindenütt-jelenvalósága. Például még egy hőszabályzó is információt jelenít meg, de tekinthetjük-e ettől tudatosnak? Legalább két lehetséges válasz van. Először, alakíthatnánk úgy az alaptörvényt, hogy csak az információk egy része bírjon tapasztalati vetülettel, mondjuk a fizikai feldolgozás függvényében. Másodszor, megengedhetjük, hogy minden információ rendelkezik tapasztalati vetülettel – ahol komplex információ-feldolgozás történik, ott komplex tapasztalat van, ahol pedig egyszerű információ-feldolgozás, ott egyszerű tapasztalat. Ha ez így áll, akkor még egy hőszabályzó-nak is lehetnek tapasztalatai, még ha sokkal egyszerűbbek lennének az elemi szintapasztalatnál is, és érzelmek és gondolatok bizonyára nem kísérnék őket. Ez első ránézésre furcsának tűnhet, de ha a tapasztalat tényleg alapvető, akkor azt várhatnánk, hogy ez a nézet széles körben elterjedt. Mindenesetre az alternatívák közötti választásnak attól kell függenie, hogy melyik építhető a leghatékonyabb elméletbe.



Az ilyen elképzelések természetesen tévesek is lehetnek. Másrészt, továbbfejlődhetnek egy hatékonyabb indítvánnyá, amely előrelátná a tudatos tapasztalat pontos struktúráját az agyban végbemenő fizikai folyamatok alapján. Ha sikerül a terv, akkor jó okunk lesz rá, hogy az elméletet elfogadjuk. Ha nem, más utakat keresünk, és talán alternatív alapelméleteket dolgozunk majd ki. Így előbb-utóbb talán megoldjuk az elme legnagyobb rejtélyét.

## Táncoló „qualia” egy szintetikus agyban

Sokan már önmagában is érdekesnek találják azt a kérdést, hogy a tudat létrejöhet-e egy komplex, szintetikus rendszerben. Még ha évtizedek vagy -századok telnek is el, amíg egy ilyen rendszer elkészül, egy egyszerű gondolatkísérlet komoly bizonyítékokkal szolgál arra, hogy egy mesterséges agy, megfelelő szerveződéssel, valóban ugyanolyan tudatos tapasztalatokkal rendelkezik, mint egy emberi lény.

Képzeljünk el egy szilícium-alapú rendszert, amelyben a *chipek* úgy vannak elrendezve és úgy működnek, mint agyunk neuronjai. Azaz, a szilícium-rendszerben minden *chip* pontosan ugyanazt teszi, mint természeti megfelelője, és ugyanolyan módon kapcsolódik a környező elemekhez. A mesterséges rendszer viselkedése így pontosan olyan lenne, mint a mienk. A döntő kérdés pedig: hozzánk hasonlóan tudatos lesz-e?

Az érvelés kedvéért tegyük fel, hogy nem. (Itt a *reductio ad absurdum* néven ismert érvelési technikát használjuk, ahol az ellentétes feltevésből kiindulva bizonyítjuk, hogy az tarthatatlan következtetésre vezet.) Azaz, vagy más tapasztalatai vannak – például a kék tapasztalata, amikor más vöröset lát –, vagy semmilyenek. Az első esetet fogjuk megvizsgálni; az érvelés hasonlóan halad mindkét esetben.

Mivel a *chipek* és a neuronok ugyanazzal a funkcióval bírnak, ezért megcserélhetők, a megfelelő érintkezési felület közbeiktatásával. Így a *chipek* helyettesíthetők a neuronokat, olyan esetek kontinuumát hozva létre, amelyekben a neuronokat fokozatosan növekvő arányban *chipekkel* helyettesítjük. A rendszer tudatos tapasztalata a kontinuumon haladva szintén változni fog. Például a látóközpontod összes neuronját helyettesíthetjük egy azonos módon szervezett szilícium-változattal. Az eredmény egy olyan, mesterséges látóközpontú agy lesz, amely az eredetiétől eltérő látási tapasztalattal fog bírni: ahol eddig vöröset láttunk, most lilát fogunk látni (esetleg halványrózsaszínt, abban az esetben, ha a tisztán szilícium rendszernek nincsenek tapasztalatai).

Ezután egy kétpozíciós kapcsolóval mindkét látóközpontot hozzákapsoljuk az agyunkhoz. A kapcsoló egyik helyzetében a természetes látóközpontunkat használjuk, míg a másikban a mesterségeset. Amikor a kapcsolót átkapsoljuk, látási tapasztalatunk vörösről lilára vált, vagy fordítva. Ha a kapcsolót ismételten kapcsolgatjuk, tapasztalataink „táncolni” kezdenek a két különböző tudati állapot (vörös és lila) között, amelyeket „qualia”-nak szokás nevezni.

Mivel viszont agyunk szerveződése nem változik, viselkedésbeli változás sem következhet be a kapcsoló átváltásakor. Következésképp, ha megkérdezzük,

hogy mit látunk, azt fogjuk válaszolni, hogy semmi sem változott. Azt fogjuk állítani, hogy vöröset látunk, és csak azt láttuk eddig is, még ha a két szín ott táncol is a szemünk előtt. Ez a következtetés annyira ésszerűtlen, hogy legjobb az eredeti feltevés *reductio ad absurdum*aként elfogadni – azaz, a neuronokból felépülő aggyal azonos szerveződésű és azonos működésű mesterséges rendszernek az agyhoz képest eltérő tudatos tapasztalatai vannak. A feltevés visszavonása az ellentétes feltevést bizonyítja: hogy az ugyanolyan szerveződésű rendszereknek ugyanolyan tudatos tapasztalattal kell rendelkezniük.

## Hozzászólás David J. Chalmers írásához Miért lehet képes az idegtudomány a tudat magyarázatára?

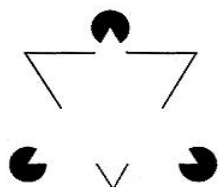
Francis Crick – Christof Koch

Úgy gondoljuk, hogy a tudat magyarázatának pillanatnyilag az a legjobb módja, ha felkutatjuk azt, amit a tudat idegi korrelátumainak nevezünk, vagyis azokat az agyi folyamatokat, amelyek a tudatért legközvetlenebbül felelősek. Azáltal, hogy az agykéregben lokalizáljuk a tudattal legjobban korrelálható idegsejteket, és rájövünk arra, hogy ezek az agyban másutt elhelyezkedő idegsejtekkel milyen kapcsolatban állnak, kulcsfontosságú bepillantást nyerhetünk abba, amit David J. Chalmers a nehéz problémának nevez: annak kimerítő magyarázatába, ahogyan az agyi folyamatokból a szubjektív tapasztalat előáll.

Elismerjük Chalmers érdemét abban, hogy a nehéz problémát már e korai lépésben bátran felismerte és vizsgálat tárgyává tette, bár nem minden gondolkísérletéért lelkesedünk. Ahogy mi látjuk, a nehéz probléma további kérdésekre bontható: Miért tapasztalunk egyáltalán bármit? Mi hozza létre a konkrét tudatos tapasztalatot (mint amilyen a kék kéksége)? Miért lehetetlen a szubjektív tapasztalat bizonyos vetületeit másokkal megosztani (azaz miért magánjellegűek)? Azt gondoljuk, hogy az utolsó kérdésre találtunk választ, illetve az első kettőre is van javaslatunk, és ezek az explicit idegi megjelenítés néven ismert jelenség köré csoportosulnak.

Mit jelent az „explicit” ebben a szöveggörnyezetben? Ezt talán egy példával határozhatnánk meg a legjobban. Egy arc képére adott válaszként a dúcsejtek az egész retinán kisüléseket okoznak, igen hasonlóan a televíziókészülék képernyőjéhez, az arc implicit megjelenítését hozván létre ezáltal. Ugyanakkor sok egyéb részletre is válaszolhatnak a képen belül, mint például az árnyékok, vonalak, egyenlőtlen megvilágítás, és így tovább. Ezzel szemben a látóközpont hierarchiájában magasan elhelyezkedő idegsejtek leginkább az arcra válaszolnak, vagy akár az egy bizonyos megvilágításban látott arcra. Ezek a neuronok segítik hozzá az agyat ahhoz, hogy az arcot explicit módon jelenítse meg. Ezen idegsejteknek agyvérzés vagy más sérülés következtében történő elvesztése prozopagnóziához vezet, ami az egyén képtelensége arra, hogy ismerős arcokat tudatosan felismerjen – a sajátját is beleértve, bár az arcot mint arcot továbbra is felismeri. Ehhez

hasonlóan, a látómező más részeinek sérülése a színlátás képességének elvesztéséhez vezethet, miközben az egyén fekete-fehér árnyalatokat továbbra is észlel, és a szem színreceptorai sincsen sérülés.



*A Kanizsa-háromszög azokat az idegsejteket ingerli, amelyek az ilyen illuzórikus körvonalakat kódolják explicit módon.*

A látási információ minden szinten újrakódolódik, legtöbbször fél-hierarchikus módon. A retina dúcsejtjei a fényes területekre válaszolnak. Az elsődleges látóközpont idegsejtjei a vonalakra vagy szélekre való válaszáadásra a legalkalmasabbak; a magasabban lévő idegsejtek a mozgó körvonalakat részesíthetik előnyben. Még főnnebb helyezkednek el azok, amelyek arcokra vagy ismerős tárgyra válaszolnak. A csúcson azok vannak, amelyek az agy premotorikus és motorikus struktúráiba vezetnek, ahová idegingerületeket továbbítanak, elindítván az olyan cselekvéseket, mint amilyen a beszéd, vagy az, hogy félrelépünk egy közeledő gépjármű elől.

Chalmers, akárcsak mi, úgy véli, hogy egy tapasztalat szubjektív vetülete szorosan össze kell hogy függjön az illető vetületnek megfelelő idegsejtek kislésével (ezek az idegi korrelátumok). Leír egy jól ismert gondolkísérletet, mely egy feltételezett tudós, Mary köré épül, aki a színlátás szakembere, de sohasem látott – színeket. Úgy gondoljuk, Mary azért nem tudja, milyenek a színek, mert agyában sohasem jött létre valamely szín explicit idegi megjelenítése, csupán a színekkel kapcsolatos szavaké és fogalmaké.

Ahhoz, hogy egy szubjektív látási tapasztalatot leírjunk, az információnak el kell jutnia az agy motorikus kimeneti szintjére, ahol is a verbalizáció vagy más cselekvések számára hozzáférhetővé válik. Ez a továbbítás mindig az információ újrakódolását jelenti, úgy, hogy a motorikus idegsejtek által kifejezett explicit információ összefügg, de nem azonos a látási hierarchia valamely szintjén a szinttapasztalathoz kapcsolódó idegsejtek által kifejezett információval.

Nem lehetséges tehát szavak vagy fogalmak útján közölni a szubjektív tapasztalat pontos milyenségét. Lehetséges viszont a szubjektív tapasztalatok – például a vörös és a narancsszín – közötti különbséget közölni. Ez azért lehetséges, mert a magasszintű látásért felelős agykéreg-területen jelentkező különbség továbbra is a motorikus szintek közötti különbséggel fog összekapcsolódni. A következmény az, hogy sohasem magyarázhatjuk meg másoknak tudatos tapasztalataink milyenségét, csak más tapasztalatokkal való viszonyát.

A két másik kérdés, arra vonatkozóan, hogy miért vannak tudatos tapasztalataink, és mi hoz létre egyes egyedi tapasztalatokat, már nehezebbnek tűnik. Chalmers azt veti fel, hogy ezek a „tapasztalatnak” mint a világ új alaptulajdon-

ságának a bevezetését kívánják, ami a szervezet információ-feldolgozási képességével függene össze. De mely idegi információk hozzák létre a tudatot? És mi feleltet meg egy bizonyos fajta információt inkább a kék kékségének, mint a zöld zöltségének? Az ilyen problémák éppoly nehéznek tűnnek, mint bármi más a tudat tanulmányozásában.

Mi egy alternatív megközelítést részesítünk előnyben, amely a „jelentés” fogalmát foglalja magában. Milyen értelemben mondhatjuk azt, hogy egy arcot explicit módon kódoló neuronok az arc jelentését közvetítik az agy többi része számára? Egy ilyen tulajdonságnak olyan szinaptikus kapcsolatok mintáit kell összefüggésbe hoznia a sejtek továbbviteli területeivel, melyek a kapcsolódó fogalmakat explicit módon kódoló idegsejtekhez vezetnek. Ezek a kapcsolatok végül a motorikus kimenethez visznek. Például az egy bizonyos arcra válaszoló neuronok összefügghetnek az illető személy nevét kifejezőkkel, illetve a hangját kifejezőkkel, a róla való emlékeket kifejezőkkel és így tovább. Az ilyen idegsejtek közötti kapcsolódásoknak viselkedési szempontból hasznosaknak, más szóval a test és a külvilág felől jövő visszacsatolással összeférhetőeknek kell lenniük.

A jelentés az agykérgi rendszer beláthatatlan asszociációs hálózatában megjelenések továbbiakkal való kapcsolódásaiból jön létre, egy szótárhoz vagy összefüggő adatbázishoz hasonlóan. Minél változatosabbak a kapcsolódások, annál gazdagabb a jelentés. Ha előző prozopagnóziás esetünkhöz hasonlóan az ilyen arc-neuronok szinaptikus kimenete elzáródik, a sejtek továbbra is válaszolnának az illető arcára, de nem lenne ezzel kapcsolatos jelentés, tehát tapasztalat is csak jóval kevesebb. Az arcot az alany látná, de mint olyat nem ismerné fel.

Neuronok csoportjai természetesen új funkciókat is kaphatnak, lehetővé téve az agy számára új kategóriák megtanulását (beleértve az arcokét), valamint új kategóriáknak meglévővel való összekapcsolását. Némely primitív kapcsolódások, mint például a fájdalom, bizonyos mértékben velünk születettek, de később, az élet során árnyaltabbakká válnak.

Lehetséges, hogy az információ valóban kulcsfogalom, ahogyan azt Chalmers gyanítja. Ahhoz, hogy erről jobban megbizonyosodjunk, meg kell majd vizsgálnunk a magas szinten párhuzamos információ-áramokat, amelyek – a neuronokhoz hasonlóan – komplex hálózatokba kapcsolódnak. Hasznos lenne meghatározni, milyen tulajdonságokkal kell egy idegi hálózatnak (vagy valamely egyéb komputációs formának) bírnia ahhoz, hogy jelentést hozzon létre. Megeshet, hogy az ilyen gyakorlatok a jelentés idegi alapja felé mutatnak majd. A tudat nehéz problémája így új megvilágításba kerülhetne. Vagy esetleg el is tűnhetne.

*Fordította Telegdi-Csetri Áron*